

NAT'L INST. OF STAND & TECH
A11106 222618

NBS
PUBLICATIONS

O

NBSIR 81-2354

Federal Requirements for a Federal Information Processing Standard Data Dictionary System

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Institute for Computer Sciences and Technology
Center for Programming Science and Technology
Washington, D.C. 20234

September 1981



QC U.S. DEPARTMENT OF COMMERCE
100 NATIONAL BUREAU OF STANDARDS
.U56
81-2354
1981
c. 2

NATIONAL BUREAU
OF STANDARDS
LIBRARY

OCT 14 1981

Not a - 3rc

Q100

USG

10.31-0354

1981

1-2

NBSIR 81-2354

**FEDERAL REQUIREMENTS FOR A FEDERAL
INFORMATION PROCESSING STANDARD DATA
DICTIONARY SYSTEM**

Patricia A. Konig
Judith J. Newton

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Institute for Computer Sciences and Technology
Center for Programming Science and Technology
Washington, D.C. 20234

September 1981

**U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, Secretary
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director**

TABLE OF CONTENTS

	Page
1. INTRODUCTION	2
1.1 Background	2
1.2 Expected Benefits	2
1.3 NBS Project Approach and Status	3
1.3.1 Project Approach	3
1.3.2 Project Status	4
1.4 Scope of Report	5
1.5 DDS Definitions and Terminology	7
2. REACTION TO THE PROSPECTUS	8
2.1 Federal Responses to the Prospectus	8
2.2 Non-Federal Responses to the Prospectus	9
3. FEDERAL AGENCY INTERVIEW SUMMARY	11
3.1 Purpose and Use of the DDS	11
3.2 DDS Operating Environment	13
3.3 Problems with DDS Use	16
3.4 Benefits Obtained	17
3.5 Trends and Future DDS Use	18
4. DDS ITEM RATING RESULTS	21
5. CONCLUSIONS	47
5.1 Specific Conclusions and Unresolved Issues ...	47
5.1.1 DDS Entities	48
5.1.2 DDS Attributes	48
5.1.3 Relationships	49
5.1.4 Operating Modes and Control Functions	50
5.1.5 Operating Modes	50
5.1.6 Control Functions	50
5.1.7 Language Functions and Format	51

5.1.8 DDS Interfaces and Functions	52
5.1.9 Standard Printed Outputs	54
5.1.10 DDS Administrator Support and Utility Functions.....	54
5.2 General Conclusions	55
APPENDIX A : QUESTIONNAIRE	57
APPENDIX B: AGENCIES AND SUBCOMPONENTS	73

PREFACE

We gratefully acknowledge the cooperation of the Federal users and developers of Data Dictionary Systems who assisted NBS in identifying current and projected Federal requirements for data dictionary software. The names of the individuals interviewed are listed in alphabetical order.

Theodore Albert	Daniel M. Breeden
Patricia Brooks	Charles Cardiff
Victor Cohen	Frank Coleman
John Cottam	Ruth Dyke
Nathaniel Gadson	Mary Kay Ganning
Gayle Gillingham	Ann Glascock
Robert Hegland	Major William Hill
Charles Honiker	Risa Katz
Wayne Koch	Robert Levi
Larry Maddox	John Martin
Michael J. Masterson	Mary Lou Melley
Steven Miers	Janice Moore
Captain Duncan Murray	Louis Pizzoli
Dwight Reynolds	Lawrence Ries
Elizabeth Scott	William Selfridge
Franklin Stack	James Stovall
Jerry Szablowsky	Frank Tagler
Robert Tiger	Jeffrey White
Major Anthony J. Winkler	.

This document identifies the data dictionary systems by trade name to portray current Federal usage of commercially available DDS's. Identification of a system does not imply a recommendation or endorsement by the National Bureau of Standards. Similarly, the omission of a system does not imply anything detrimental.

Several members of the Data Management and Programming Languages Division of NBS participated in interviews or read and criticized early versions of this report. They are Dr. Dennis W. Fife, Division Chief; Roy Saltman, Group Leader, Belkis Leong-Hong and Dr. Alan Goldfine. Comments and suggestions should be addressed to:

Mrs. Patricia A. Konig
Data Management and Programming
Languages Division
National Bureau of Standards
Washington, D.C. 20234

FEDERAL REQUIREMENTS FOR A FEDERAL INFORMATION
PROCESSING STANDARD DATA DICTIONARY SYSTEM

Patricia A. Konig
Judith J. Newton

This report presents information and preliminary conclusions about Federal agencies' requirements for a Federal Information Processing Standard Data Dictionary System. Some initial requirements were identified through analysis of comments made on the "Prospectus for Data Dictionary System Standard," NBSIR 80-2115, an earlier product which describes NBS efforts to develop a standard. Most of the data used to develop preliminary conclusions on Federal requirements was collected during interviews with Federal Government users and developers of data dictionary systems. Comments received on the Prospectus and data collected during the interviews are summarized. Preliminary conclusions and issues being investigated also are presented.

Key words: Computer program; data dictionary system; data inventory; data management; data standards; database; database management system; documentation; Federal Information Processing Standards Publication; requirements; software.

1. INTRODUCTION

1.1 Background

The Institute for Computer Sciences and Technology of the National Bureau of Standards (NBS-ICST) has planned a Data Dictionary System (DDS) standard as an objective of the Federal Information Processing Standard (FIPS) program. The FIPS program derives its authority from Public Law 89-306 (The Brooks Act) and from Executive Order 11717. Under the latter, the Secretary of Commerce has final authority to approve Federal Government data processing standards.

The Data Dictionary System standard is a planned product in a family of standards and guidelines for data management software and practices that ICST is developing. The FIPS DDS will be a software specification which Federal agencies may use for procurement purposes in conjunction with Federal Property Management Regulations (FPMR). The standard would not require an agency to use a data dictionary or to use one in a prescribed manner.

1.2 Expected Benefits

A DDS can provide the following benefits:

1. Better control and management of an agency's information resources, through improved (i.e., centralized, rigorous, and standardized) data definitions, and better data collection and handling procedures;
2. Increased security and access control for the database environment;
3. Effective aid to software development, modification, and maintenance throughout the system development life cycle; and
4. Improved documentation for databases, programs, and systems.

A Federal Information Processing Standard for a Data Dictionary System will provide additional benefits. It will contain standard specifications which can be used in the selection, evaluation, and procurement of DDS software. The FIPS DDS will aid in the portability of software and related data. Portability is the ability to transfer data, including the DDS contents, from one computer system to another,

without an agency being required to:

- o Re-create or re-enter data descriptions, except by an unload/reload process; or
- o Modify significantly the DDS application that is being transported.

The FIPS DDS also will support portability of acquired skills. Agency personnel will not need additional training to learn new user languages in order to use another DDS implementation.

1.3 NBS Project Approach and Status

1.3.1 Project Approach. The objective of the NBS project is to develop a standard that will support Federal agency requirements, and that also can be implemented by a wide spectrum of software suppliers. The project, which was initiated in late fiscal year 1979, is based on the following approach:

1. Close and continuing interaction with the Federal community to determine if:
 - o a particular capability is required by a sufficiently large segment of the Federal community;
 - o Federal users plan to, or would like to, incorporate the capability in their data processing operations; and
 - o it is desirable to have the capability in view of known constraints.
2. In-depth technological assessments and intensive consultation with hardware and software vendors, the research community, and Federal developers of in-house data dictionary systems, to determine:
 - o whether it is technologically practical to develop a particular capability in the near future, e. g. next 3 to 5 years; and
 - o if technically feasible, whether it is economical for the software industry to produce such a capability in a competitive market.

3. Periodic reports of current NBS plans;
4. Formal solicitation of comments and suggestions from all affected communities throughout the entire developmental and standardization process; and
5. Continuing interchange with the American National Standards Institute (ANSI) Technical Committee, X3H4, on Information Resource Dictionary System (IRDS).

1.3.2 Project Status. The work plan for developing the FIPS DDS is divided into the following four phases:

1. State-of-the-art assessment of DDS technology;
2. Requirements Definition;
3. DDS Functional Specification Development; and
4. Standard Specification.

During the first phase, now completed, relevant literature and existing commercial and Federally-developed data dictionary systems were analyzed. Features and capabilities in the current generation of DDS's were identified. A preliminary assessment identified projected technological trends and issues that warranted further investigation. The following two products were published during the first phase:

1. "Prospectus for Data Dictionary System Standard," NBSIR 80-2115. The Prospectus discusses the use of data dictionaries, and describes NBS efforts to develop a Federal Information Processing Standard for Data Dictionary Systems. ICST encourages technical input regarding appropriate content for a FIPS DDS.
2. "Guideline for Planning and Using a Data Dictionary System," Federal Information Processing Standard Publication (FIPS PUB) 76. This publication discusses the capabilities and uses of data dictionary systems. It also provides Federal agencies with guidance on DDS selection, planning for use of a DDS, DDS implementation, and operational usage of a DDS.

The second phase of the project is currently in progress. Interviews have been conducted recently with Federal agencies to identify current and projected requirements for data dictionary software. Interview results, as well as comments received on the Prospectus, are summarized in this

report. The report will be disseminated to all Federal agencies, DDS vendors, and other individuals and organizations working with data dictionaries. All comments received will be reviewed and evaluated as the project proceeds.

Using the results of the first two phases, ICST will work closely with the ANSI-X3H4 technical committee and with nationally recognized experts on data dictionary systems to develop DDS Functional Specifications in the third phase. The DDS Functional Specifications are scheduled for publication in late fiscal year 1982.

The focus of the final phase will be to develop the candidate FIPS DDS. Current plans are to publish the candidate standard in fiscal year 1983.

1.4 Scope of Report

Readers of this report are presumed to be familiar with general data processing concepts and with the general concepts and purpose of a Data Dictionary System. Readers are referred to the Prospectus for an overview of the concepts, purpose, and capabilities of DDS's.

In the Fall of 1980, the Prospectus was disseminated to all of the Federal senior management officials. Federal agencies appointed these officials in 1979, at the request of the Office of Management and Budget, to provide ICST with agency views on ADP concerns and standards needs. The Prospectus was distributed also to data dictionary vendors and to organizations expressing an interest in the NBS program. Responses that have been received on the Prospectus are summarized in Chapter 2.

The results of Federal agency interviews are presented in Chapters 3 and 4. Three criteria were used to select the agencies to be interviewed. First, interviews were limited to the vicinity of Washington, D. C. because of travel limitations and time constraints. Second, interviews were conducted only with agencies which have experience in the implementation or use of data dictionary systems. Third, agencies were selected to provide a balance between large and small, and between civil and defense agencies. Based on these criteria, interviews were conducted in the following agencies (for a detailed list of the agency subcomponents, please see Appendix B):

- o Air Force Data Service Center
- o Defense Communications Agency
- o Office of the Assistant Secretary of Defense (Comptroller)
- o Department of Labor, Employment and Training Administration
- o Department of Labor, Office of the Assistant Secretary for Administration and Management
- o Environmental Protection Agency
- o Internal Revenue Service
- o Library of Congress
- o National Archives and Records Service
- o Naval Data Automation Command, Navy Regional Data Automation Center
- o Small Business Administration
- o Social Security Administration
- o Geological Survey
- o Veterans Administration

Although fourteen interviews were conducted, two of them addressed the same DDS implementation. (National Archives and Record Service had an Inter-Agency Agreement with the Department of Labor, Employment and Training Administration (ETA) to provide assistance in the selection and implementation of ETA's DDS.) Thus, the use of data dictionary systems in thirteen agencies was reviewed. From one to five agency representatives participated in each interview for a total of thirty-seven interviewees whose names appear in the Preface.

Interviewees' responses to the questions represent their best judgments, and not an official agency assessment. All interviews were conducted using the same two data collection instruments. One of these was an Interview Guide. Questions asked during each interview were based on the Interview Guide. These questions covered: (1) discussion of the project approach; (2) questions pertaining to an agency's experience using a data dictionary; and (3) solicitation of their views on the future trends in the use and capabilities of data dictionaries. Responses to these questions are summarized in Chapter 3.

The other data collection instrument was the Rating Form which lists 96 generic features of existing and conceptual data dictionary systems. This instrument was sent to agency personnel in advance of the interview. They were asked to rate each item on a scale of 1 to 5 which ranged from "not required" to "indispensable." Several agencies completed more than one Rating Form. Rating results for the eighteen Rating Forms that NBS received appear in Chapter 4.

Some conclusions have been reached based on comments received on the Prospectus, the Item Rating Form results, and the interview question responses. These conclusions, subject to additional review through workshops and reactions to this report, are discussed in Chapter 5. Technical issues requiring further analyses are also presented.

1.5 DDS Definitions and Terminology

This section contains working definitions for terms which are used in the remainder of this report.

Attribute - a property or characteristic of an entity.

Data Dictionary System (DDS) - a computer software system that provides for recording, storing, and processing information about an organization's significant data and data processing entities.

DDS Entry - the conceptual grouping of an entity and all its associated attributes, which is entered in the DDS. The DDS is thus a collection of entries.

Entity - any named concept, object, person, event, process, or thing that is the subject of stored or collected data.

Entity Type - a class of entities having the same attributes. The class of "data element" is an example of one entity type.

Extensibility - the ability to extend the original range of entity types, attributes and/or relationships of a DDS to include those unique to any one user's environment.

Occurrence - a specific instance or value of an attribute, such as "Test" Status. The occurrence appears within a DDS Entry.

2. REACTION TO THE PROSPECTUS

The first section of this Chapter summarizes the comments received to date from Federal agencies. Comments received from non-Federal sources appear in the second section. Reaction to the Prospectus was voluntary and not limited to a pre-defined time period. Comments received after April, 1981 will be summarized in future reports.

2.1 Federal Responses to the Prospectus

The following agencies sent formal comments on the Prospectus through their senior management officials:

- o Department of Energy
- o Federal Communications Commission
- o National Aeronautics and Space Administration
- o Nuclear Regulatory Commission
- o Tennessee Valley Authority
- o Department of Agriculture
- o Department of Justice
- o Veterans Administration

Several other agencies submitted informal comments.

The point emphasized most frequently was that the FIPS DDS must be "flexible". Several agencies want a data dictionary system that will be usable in a variety of hardware and software environments. One agency stated that the DDS must support a distributed environment and that it would be desirable to have a DDS that could operate on a minicomputer. Most of the agencies addressing the issue of flexibility requested that ICST consider developing a "family" of standards. There were concerns that the concept discussed in the Prospectus of specifying one standard, which included a core module and additional optional modules, might not provide the needed flexibility for the Federal community. There were also expressions of concern that ONE standard might result in a DDS that would be either too simple or too sophisticated and costly. Some agencies added that the FIPS DDS should not preclude the implementation of the DDS as an application of a Database Management System (DBMS) or a file management system.

Most agencies expressed a requirement for a DDS that would be independent of any specific DBMS. But at the same time, they felt the DDS should have the capability to support, to some degree, different DBMS's. An often-expressed requirement was the need for the DDS to interface with COBOL, FORTRAN, and other problem-oriented and scientific-based languages.

Other agencies felt that the Prospectus did not address certain DDS capabilities in sufficient depth. For example, several agencies expressed the requirement for access and security controls. One agency required access restriction at the individual entity level. The DDS user language(s) was cited as another important aspect that needed further elaboration. In particular, ease of understanding and ease of use were areas of concern. Another respondent stated that relationships among data entities needed to be more clearly and precisely defined.

The remainder of the comments addressed individual agency needs. The following were included:

- o It should be possible to document more than "data" descriptions. It is considered a requirement or at least "highly desirable" to be able to describe the functions of an organization.
- o It would be desirable to have an option to include "unlimited comments" (i.e., no restriction on the size of the "description" field) for each entity type. This would provide the ability to capture pertinent information that otherwise cannot be specified.
- o It should be possible to specify the owner and/or the person or office that is responsible for maintenance of an entity or entities.

2.2 Non-Federal Responses to the Prospectus

Several organizations which currently market data dictionary software or are in the process of developing data dictionary software commented on the Prospectus. Comments were received also from a university and from a large private organization, both working with data dictionaries.

The comments from vendors focused primarily on the working assumption that the FIPS DDS would specify a stand-alone DDS that was independent of specific vendor hardware or software. Vendors were concerned that elimination of a dependent DDS may degrade efficiency and increase resource requirements. They also stated that many of the benefits of a DDS are derived from its full integration and participation with other software systems such as a DBMS. One major advantage cited for integration with a DBMS is that users can employ the same query and reporting protocols for access to both data in the DBMS and its description in the DDS.

Vendors indicated that they must maintain a high degree of consistency and continuity in upgrading to new technology. They recommended that the standard specify language functions and capabilities but not the precise syntax. Likewise, an acceptable level of DDS overhead should be estimated, but the method of implementation should be a vendor option.

Several vendors felt that the term "Data Dictionary" was too limiting, and recommended that another name be selected that would better convey the concept of management of both manual and computer resources. Another recommendation focused on the desirability of acknowledging the user environment in both the Functional Specifications and the Federal Information Processing Standard. Such factors as ease of use, interactive versus batch access, multi-users and concurrency controls, customization of reports, and end-user access facilities should be addressed, according to these comments.

Responses from the university and the private organization strongly supported an extensibility capability which would enable users to define their own entities, attributes, and relationships. They also addressed technical issues and trends which they felt needed to be reviewed during the development of the standard. These issues included the interface, or integration, of the DDS with a DBMS; partitioning and allocation of control among the DDS, DBMS, and language processors; and simultaneous use of the DDS by diverse applications.

A number of comments concerned the capabilities that were emphasized in the Prospectus. Certain respondents felt that the use of a DDS as a "data describer" was emphasized to the detriment of its use in the analysis and design of applications. It was also stated that version control and active control characteristics, such as control over data structures and control over the computer programs that process specific data structures, should be reviewed in greater depth.

Several approaches will be used to review the comments and resolve the issues and conflicting viewpoints included in this Chapter. Respondents will be invited to the workshops scheduled to occur during the next year. Federal requirements and ICST's proposed solutions to technical and economic issues identified in this and subsequent Chapters will be presented. Workshop participants' reactions and input to the content and scope of the FIPS DDS will be encouraged. Interviews with the DDS vendors and additional discussions with Federal agencies and others knowledgeable in the implementation and use of DDS's also are planned.

3. FEDERAL AGENCY INTERVIEW SUMMARY

During the Federal agency interviews, data was collected about how the DDS is or will be used. This data was used to clarify interview responses, interpret rating results, and develop conclusions on Federal requirements. For example, some agencies are using their DDS solely to inventory data elements and forms in order to standardize the data elements and eliminate redundant data collection. These agencies generally gave low ratings to items associated with the analysis, design and maintenance of ADP systems. ICST will use interview results during the next phase of the project to develop DDS Functional Specifications that support existing and planned operating requirements.

The Federal Agency Interview Guide was divided into six sections. The complete Interview Guide is included in Appendix A. Part I solicited information about the interviewees' involvement with a data dictionary system and their position in the organizational structure. This information appears in Appendix C.

Part II was designed to clarify any questions that interviewees had about the ICST project or about the items or terminology in the Rating Form. Chapter 4 contains the comments which were made on the Rating Form. These comments appear after the statistical summary of the ratings for each item.

Parts III to V of the Federal Agency Interview Guide addressed the current use of the DDS. Questions focused on the (1) Purpose and Use of the DDS; (2) DDS Operating Environment; (3) Problems with DDS Use; and (4) Benefits Obtained. Part VI contained questions on anticipated trends and future plans for DDS use. Responses to questions in Parts III to VI are summarized in this Chapter.

3.1 Purpose and Use of the DDS

Most of the interview questions were designed to obtain experiential data on the current use of data dictionary systems in the Federal government. In answer to why a DDS was obtained, one interviewee described a situation that exists to some degree in the other agencies. "Literally thousands of computer programs have been developed over many years. There has been little, if any, effort to impose naming standards. In the few areas where the importance of standard names and descriptions was recognized and enforced, the

standards which evolved were limited to specific program areas. We now find that analysts from one program area cannot easily communicate their ideas, concepts, or data requirements to their counterparts in another program area. For example, the term 'individual' has one meaning in one system and another meaning in a different system."

Generally, the data dictionary serves as a central reference file which documents existing data. One agency obtained their DDS to assist in a large conversion effort to identify where and how data elements were used. Several agencies use the DDS as a tool to eliminate or minimize the collection of redundant data. The DDS is checked to determine if the needed data exists and if so where it is located. To help eliminate redundant data collection, two agencies use the DDS to identify forms which are used in collecting information from the public or from other Federal agencies and to describe the data elements which appear on these forms. Other agencies use the DDS to eliminate duplication of effort during requirements analysis by identifying existing data which can satisfy the new requirements.

Agencies also use the DDS as a tool for data element standardization. Keyword searches and character-string analysis facilities are some of the capabilities which help identify similar or redundant data. One agency developed a thesaurus to use in conjunction with its DDS. Terms from the thesaurus, which classify the type of data, are entered as keywords in the DDS. The DDS, by serving as a central reference file for one or more program areas, increases the awareness of those standards that are developed.

Several interviewees' DDS's are used to assist in the design of application software and databases. One uses the DDS as a "scratch pad" for proposed applications. Others use it to prepare COBOL data division statements, PL/1 structures, and DBMS schema entries. None of the DDS's exercise total, active control over other systems. In the four agencies where the DDS exerts some degree of control, two generate source data for application programs through COPY libraries. One agency generates schemas and subschemas for its DBMS, but refines them manually before actual use. Another agency accesses the DBMS through the DDS, but the DBMS can also be reached by other methods. In only one instance is the DDS used to enforce the use of standards. This is accomplished by DBMS edit routines which call the DDS to check object code for standard data elements. Deviations from standard descriptions are then printed.

The DDS is being used by ten agencies to prepare record specifications and documentation for application programs and logical database design. In one agency, record specifications have been standardized through use of the DDS. The DDS, in this case, helps increase productivity by eliminating the need to document record specifications separately. It also facilitates management review of the system development process because the specifications are always presented in the same format, and thus discrepancies are more readily identifiable.

Others use or will soon use the DDS to:

- o Help identify the costs and benefits of information resource management
- o Describe network topology and the location of data in the network
- o Index scientific data and its location.

The DDS, by documenting users of data, also helps prevent one person from making a change in data or record descriptions that could affect other people.

None of the DDS's are currently used to generate test data or to control operations and scheduling. One agency attempted to use its DDS to control operations and scheduling, but synchronization of the DDS and the DBMS became a major problem. The agency is now using other software for this purpose. Another agency, however, is using the DDS to help its personnel establish priorities for system operations.

3.2 DDS Operating Environment

Agencies were asked a series of questions about the existing DDS operating environment. Responses on the processing environment, including DDS-DBMS relationship and DDS-programming language interface, are summarized in Figure 1. Other questions focused on the number and size of DDS's in the agency, and DDS operating characteristics such as response time and language.

As shown in Figure 1, some agencies use only one DDS. Others have several DDS's. One agency, for example, has 9 separate dictionaries, all of which are processed by the same software. Seven of these DDS's contain program-specific data for different organizational components. One DDS documents an integrated database that processes data for three of these organizational components. The ninth DDS

DATA DICTIONARY SYSTEM PROCESSING ENVIRONMENTS				
DDS SYSTEM	COMPUTER TYPE/ OPERATING SYSTEM	DBMS USED	DBMS-DDS RELATIONSHIP	PROGRAMMING LANGUAGES USED WITH DDS
ADABAS	IBM 360/65 OS MBT TSO	ADABAS	DDS INTEGRATED WITH DBMS	COBOL
DATA CATALOG 2	IBM 168/3033 MP JES 2	IDMS SYSTEM 2000	NONE	COBOL, PL/1
DATA CATALOG 2	Univac 1100/81 EXEC-8	DMS-1100	DDS generates layouts and schemas for DMS-1100	COBOL
DATA CATALOG 2 IDD-Cullinane	IBM 3033 OS, TSO, JES	NONE	N/A	NONE
DATAMANAGER	ITEL AS-6 MVS	NONE	N/A	ASSEMBLY NOW, COBOL IN FUTURE
DATAMANAGER	IBM 4331 CMS IBM 4341 MVT ITEL AS5 OS/MVS	TOTAL	DDS produces data descrip- tions for TOTAL application programs	NONE
DATAMANAGER SSI-DED*	IBM 360/168 MVS DMS-1100 Univac IMS 8	TOTAL SYSTEM 2000	NONE	COBOL, ALC
IBM DB/DC Data Dictionary	Amdahl 470 V-6 IBM 370/3033 MVS JES2	IMS/DLL Logic Library MUMS (both in-house)	REFERENCE ONLY	PL/1, COBOL, ALC
PRIDE/LOGIK	Honeywell 6880 MULTICS	Multics Relational Data Store	NONE	COBOL, PL/1
ESIS*	Amdahl V-7 MVT	Model 204 System 2000 IDMS - MRDS	ESIS- DM-4 application	NONE
IRCAS*	DEC-20	NONE	N/A	COBOL
RAS*	Univac 1100 EXEC-8 Honeywell 6000 GCOS IBM 360/370 CDC 7600	DMS 1100 WWDMDS ADABAS	REFERENCE ONLY	COBOL
VADD* HDDS* (not yet implemented)	IBM 360/65 OS (for VADD) Honeywell 6600 (for HDDS)	DM-4 TOTAL IDMS	VADD-NONE HDDS- DM-4 application	COBOL

* Federally-developed system.

Figure 1. Data Dictionary System Processing Environments

contains skeletal, management-oriented descriptions of the data contained in the other eight. Several other agencies have separate DDS's which are processed by different software. When multiple dictionaries are being used, there is currently no direct communication between them.

Six of the agencies interviewed are implementing a DDS or have used their dictionary only for a relatively short period of time. As a result, they have small dictionaries which range in size from 70 to 200 entries. In most of these DDS's, only a portion of the available entity types are being used, but most agencies plan to use additional entity types in the future. The larger dictionaries generally have 5,000 to 10,000 entries. In two of the agencies where multiple dictionaries exist, the total number of DDS entries is greater than 10,000.

Estimates of the number of DDS users range from four to 2,000 (for a system with multiple installations). The number of users is largely influenced by the length of time the DDS has been operational and by the number of organizational program areas supported. At present, ADP personnel are primarily responsible for DDS update and retrieval. In many agencies, however, DDS reports are being used by people with no ADP background to identify and eliminate redundancy and standardize data elements.

Most users are satisfied with the response time and level of system overhead associated with their DDS. One complained of slow turn-around time, but blamed the overloaded time-sharing system rather than the DDS. Another agency is unhappy with the overhead, but is interactively using a system designed to operate in batch mode. Two agencies complained that some queries take too long to process, e.g., an hour at a cost of \$500.

All systems in the agencies interviewed have security controls, although the type of control varies. The most common control involves user passwords, to the system as a whole and to levels of entity type. Other types of security involve user access control (read-only, update, and maintenance) and access by ownership of data. Integrity controls include audit trails of changes, editing of DDS input data, creation of separate test versions, and backup capabilities through both the DDS and the computer's operating system.

Agencies were asked if they felt it was important for the DDS and DBMS languages to be similar. In four agencies, the DDS language is similar to the DBMS data definition language, either because it was so designed or because the DDS is integrated with the DBMS. Three of these agencies think the two languages should be similar; the fourth

attaches no importance to this factor. The remaining agencies which use a DBMS have different DDS and DBMS languages. Only one agency which has different languages thinks that they should be similar. Others do not feel that it matters.

Interviewees also were asked if they had ever transferred data from one DDS to another or converted DDS software from one operating environment to another. Only two have been involved in a conversion. One agency converted data from an internally-developed system to a newly-purchased commercial system. Although ultimately successful, the effort was plagued by technical failures in translating the file structures. The other agency is currently testing DDS software which has been converted from a DEC-20 to a smaller computer. Preliminary tests show an increase in response time of 7-20 times and an increase in sort time from 30 minutes to 8 hours.

3.3 Problems with DDS Use

Several interviewees feel that their DDS is difficult to use. One especially mentioned the long learning curve that new users experience. Three think that the scope of available entities and attributes is inadequate, and complained about the lack of a capability for indicating relationship between entities on the same level. Another cited a weakness in the area of synonyms and cross-referencing. The 32-character limit one system imposes on entity names and the disallowance of duplicate entity names were also mentioned as problems.

The most frequently cited problems are associated with report and query capabilities. Two specific problems cited are the inability to: (1) customize ad-hoc report formats; and (2) specify the range of data to be retrieved so that all entities which contain a set of characters in the name or definition, e.g "VET," are returned. One agency manually reformats some of the standard fixed-format reports--"a laborious process."

Four interviewees have experienced problems in the processing functions of their DDS's. In one system, the update process involves switching modes from interactive to batch to verify the information. In another, backup and restore functions are expensive and difficult to use. Several agencies feel they need a better pre-processor for data input and edit. Other complaints focused on a lack of integration with user application programs and the DBMS. In some cases, significant personnel resources have been used to develop programs to test and validate application data.

Other issues concerning the use of a DDS which were raised in the course of the interviews involved:

- o Lack of available commercial packages for non-IBM equipment;
- o Organizational issues such as minimal high-level management support and resistance to change from potential users of the DDS;
- o Delayed delivery of promised interfaces and facilities from the vendor;
- o Poorly defined procedures for input control, update frequency, and system usage; and
- o Lack of an explicit methodology for determining what data should be in the DDS.

Although these topics are beyond the scope of this report, difficulty experienced with any one of them could jeopardize the effectiveness of a DDS implementation.

3.4 Benefits Obtained

Many different kinds of benefits have been realized by DDS users. One of the most often mentioned is reduction of data redundancy by elimination of identical or similar data elements. One user noted that, without a DDS, this would have been an impossible task for his agency to have done manually since over 4,000 data elements were reviewed.

Data standardization is another function aided by the DDS. One agency feels that DDS reports are a first step in creating awareness of the need for common data definitions. Another has used the DDS to identify areas where standards and improved validation procedures are necessary.

Other benefits obtained are:

- o Improved documentation of ADP applications. One agency, as an example, standardized the format of record descriptions using the DDS.
- o Increased awareness of the impact of system modifications.

- Improved reaction time in responding to new or changed requirements. Several agencies cited generation of data descriptions for new programs as a benefit because existing, standard descriptions frequently can be used, thereby saving time. Another agency uses the DDS to identify programs affected by a change in requirements and then to generate new PL/I structures for these programs.
- Improved responsiveness in answering management questions about the existence or location of specific types of data.

3.5 Trends and Future DDS Use

Agencies were asked if they planned to obtain another data dictionary system in the near future. Several agencies plan to procure new hardware which may force them to use another DDS. In the majority of cases, however, agencies plan to continue to use or to enhance their existing DDS. Planned enhancements include the following:

- Improvement of the reporting capability to permit report customization;
- Addition of DBMS and COBOL data description generation facilities;
- Use of extensibility features; and
- Improvement of the interactive update and query facility to make the DDS easier to use and to add additional retrieval capabilities, and a facility to keep track of on-line updates.

When asked about the trends they perceived in the use of data dictionaries, all agencies predicted that DDS usage will increase. All responded that the most important use of the DDS will continue to be for standardizing and controlling data. Several mentioned use of the DDS in promoting the exchange and sharing of data. Other agencies plan to continue or begin to use the DDS for: documentation, generation of COBOL and DBMS data definitions, and change-impact analysis.

On the question of the future use of multiple or centralized DDS's, however, there is no clear trend. Five of the agencies foresee centralized DDS's; four envision multiple DDS's; and four others favor multiple DDS's under the control of one centralized DDS. Agencies which plan to have multiple DDS's under the control of one centralized DDS

state that the centralized DDS will be oriented toward agency-wide information resource management rather than toward program management.

Seven of the agencies thought that the DBMS and the DDS should be closely related. Four thought that they should be independent of each other. Two stated that it should be possible to have the DDS both independent and closely related. One of these two interviewees commented that with a centralized DDS controlling several distributed DDS's, the central DDS should be independent while the decentralized DDS's should be closely associated to or should drive a DBMS.

There is a definite consensus on the future need for a DDS in an environment of distributed processing and storage. Future plans for one agency involve development of a network with the DDS an integral part of the system. Another agency plans to expand its networking capabilities and use the DDS to assist in the control of data in the distributed environment. Even those agencies that have no plans to become involved with network systems agreed that this capability will be important eventually. One interviewee stated that distributed processing and storage "...will be increasingly the environment in which a DDS must operate."

On the question of future DDS use with minicomputers, the interviewees were divided evenly. Five thought they would require a system, five did not. The others were unsure.

Most agencies were not concerned about the scope and complexity of DDS software. Of the few who expressed concern, one said that the FIPS DDS should not specify a DDS that would be too large for minicomputer implementation. Another cautioned that an overly complex FIPS might adversely affect efficiency or vendor support.

Interviewees were asked if they felt guidelines on DDS usage were more important than a FIPS. Guidelines were preferred in only four instances. In general, responses were very much in favor of both. Those who prefer a guideline feel that software technology changes too fast and that DDS technology, in particular, is too new for a FIPS. On the other hand, one agency which feels a FIPS is necessary said it is four years overdue. Interviewees who strongly support the development of a FIPS gave the following reasons:

- o A FIPS will facilitate the consistent description of data and, therefore, will help promote the exchange of data between organizational units and between agencies.

- o A FIPS is needed so that agencies can specify their DDS software requirements in standard and consistent terminology.

There was a wide spectrum of responses to the question about how an agency will interface with the Federal Information Locator System (FILS) which has been mandated under the Paperwork Reduction Act of 1980. Some of the interviewees' organizational units have worked closely with the FILS committee. Some were not aware of the FILS project. Others said that they have no reason to be involved because they are not concerned with public use forms and reports. Those who foresee an interface with FILS feel that it will mean a significant increase in the use of their internal DDS.

4. DDS ITEM RATING RESULTS

The Rating Form listed major entity types, capabilities, and features tentatively being considered for inclusion in the FIPS DDS. The ratings reflect the interviewees' best judgment of their agency's near future requirements (1981-1986) for data dictionary software. Respondents were requested to place a number from 1 to 5 in a blank next to each item to indicate the degree to which they felt it was required in their organization. The following scale was specified:

1. Item is not required in my organization.
2. Useful, but low priority.
3. Required by a subset of DDS users or in limited applications in my organization.
4. Required widely.
5. Indispensable; must be in my DDS.

Interviewees were encouraged to write comments on the Form regarding any items they felt were missing or unclear. Time was allocated during each interview to discuss these comments.

A total of thirty-seven agency representatives participated in the fourteen interviews. Several agencies completed more than one Rating Form to reflect the interviewees' different perspectives. For example, in one agency an interviewee whose responsibilities include ADP management and long-range planning completed one Rating Form. Another Form was prepared by an interviewee who is concerned with day-to-day ADP operations. The different perspectives were reflected in the ratings. As a result of situations such as this, NBS received a total of eighteen Rating Forms.

Items on the Rating Form are organized in the following manner. Section I lists items which pertain to DDS entities and attributes. Section II contains items about different types of relationships between entities that agencies might require. Items pertaining to different operating modes and control functions of a DDS appear in Section III. Items about DDS language capabilities are in Section IV. Section V contains items relating to interfaces between the DDS and other generalized or specialized software. Possible FIPS DDS reports appear in Section VI. The final section (VII) contains items pertaining to required utilities or DDS Administrator support functions.

In the remainder of this Chapter, ratings are summarized on an item by item basis. Comments that interviewees added to the Rating Form as well as comments or questions that were raised during the interview follow the item or section to which they pertain.

I. ENTITIES and ATTRIBUTES

A. ENTITIES

Information Entities:

1. "DATA ELEMENT" A named logical unit of data.
 :Indispensable - 18
2. "DATA GROUP" A set of data elements that may be referenced as a unit or by its individual elements, but to be meaningful is normally processed as a unit.
 :Indispensable - 16
 :Required widely - 1
 :Useful, but low priority - 1
3. "RECORD" A set of data elements and data groups with a logical relationship to each other.
 :Indispensable - 17
 :Required widely - 1
4. "FILE" A collection of records on either human or machine-readable media.
 :Indispensable - 17
 :Required widely - 1
5. "DATABASE" A data collection so organized for computer processing as to reduce duplicative storage and improve the independence of the stored data structure from the processing programs.
 :Indispensable - 15
 :Required widely - 1
 :Limited application - 2

6. "SCHEMA" The logical and physical description of a database that is processed and stored by a Database Management System.

:Indispensable - 6
:Required widely - 5
:Limited application - 6
:Not required - 1

7. "FORM" A frame or outline, printed on hard copy or otherwise displayed, to facilitate the desired placement of data.

:Indispensable - 9
:Required widely - 4
:Limited application - 1
:Useful, but low priority - 3
:No rating - 1

:One agency is considering using the RECORD entity to describe the FORM and the REPORT entity.

8. "REPORT" An output from an information-handling process.

:Indispensable - 9
:Required widely - 5
:Limited application - 1
:Useful, but low priority - 2
:No rating - 1

9. "DOCUMENT" A data medium and the data recorded on it, that generally has permanence, and that can be read by man or machine.

:Indispensable - 5
:Required widely - 3
:Limited application - 2
:Useful, but low priority - 3
:Not required - 2
:No rating - 3

:One interviewee stated that this item and FORM were equivalent and responded by rating FORM and REPORT Indispensable and not rating DOCUMENT. Another interviewee noted that FORM and REPORT can be omitted if DOCUMENT contained a Type attribute, but rated all three entities

Indispensable. A third interviewee considered FORM, REPORT and DOCUMENT to be beyond the scope of the implementation of the DDS at that agency, and did not rate them.

Process Entities:

Entities which describe systems and their components, including the hardware, system software, DBMS, utilities, and manual procedures.

10. "SYSTEM" A collection of people, machines and procedures organized to accomplish a set of specific functions.

:Indispensable - 15
:Useful, but low priority - 3

:One interviewee, while rating this entity Indispensable, commented that her agency divides SYSTEM into two classes: applications (such as a budget system) and facility (such as a DBMS).

11. "PROGRAM SUBSYSTEM" A collection of related programs.

:Indispensable - 10
:Required widely - 2
:Limited application - 2
:Useful, but low priority - 4

12. "PROGRAM (SUBPROGRAM, MODULE)" A series of instructions or statements that specifies actions that may or may not be taken, expressed in a form suitable for execution by a computer.

:Indispensable - 11
:Required widely - 3
:Limited application - 1
:Useful, but low priority - 3

13. "MANUAL PROCESS" Data entry, validation, manipulation of forms, or other procedure performed by one or more persons.

:Indispensable - 5
:Required widely - 2

:Limited application - 4
:Useful, but low priority - 5
:Not required - 2

Other Entities:

14. "USER" A person, organization, or group of people who use and are concerned with entities described in the DDS.

:Indispensable - 7
:Required widely - 5
:Limited application - 3
:Useful, but low priority - 2
:Not required - 1

15. "USER DEFINED ENTITY (EXTENSIBILITY)" This item allows the list of entities in a DDS to be extended to include entities unique to a given user's programmatic responsibilities.

:Indispensable - 6
:Required widely - 5
:Limited application - 5
Useful, but low priority - 2

:One interviewee noted that extensibility must be controlled at some central point.

ADDITIONAL NOTES ON ENTITIES

1. One agency needs a JOB STREAM entity to tie all the program entities together.
2. Two interviewees thought that there should be more entities concerning manual systems.
3. One interviewee expressed a need for items relating to subschemas, and a NODE entity, which could be used to define and document network topology.
4. PROCEDURE, STEP and MODULE were noted as missing but needed by one agency. It was also pointed out that problems of redundancy can occur when coding data element and data group entries from a matrix form.

B. ATTRIBUTES

IDENTIFIER ATTRIBUTES, ALIASES, AND DESCRIPTIVE ATTRIBUTES

Identifier Attributes

16. "DDS NAME" The unique identifier of the entity within the DDS. With the Version Identifier, the DDS Entry is uniquely identified. Can apply to all entities.

:Indispensable - 17
:Required widely - 1

:One interviewee could foresee a need for several entities of different types with the same name, such as records and files, or database and schema.
17. "VERSION IDENTIFIER" Distinguishes the versions of an entity from one another and, combined with the DDS Name, forms a unique key for DDS Entries. It may be either a number or a descriptive name. Can apply to all entities.

:Indispensable - 15
:Required widely - 1
:Limited application - 1
:Useful, but low priority - 1
18. "STATUS" Position of the entity in its life cycle. Examples could be "Test," Production," "Archival," etc. There can be several versions in each status category. Can apply to all entities.

:Indispensable - 11
:Required widely - 5
:Limited application - 1
:No Rating - 1
19. "IDENTIFICATION SYMBOL" A set of characters assigned by an agency or the Office of Management and Budget in order to classify or control a form, report or document entity.

:Indispensable - 9
:Required widely - 2
:Limited application - 1
:Useful, but low priority - 1
:Not required - 2
:No rating - 3

20. "ALIASES" The set of alternative names of an entity as used outside the DDS. These may include common names, synonyms, COBOL names, PL/I names, etc. Can apply to all entities.

:Indispensable - 17
:Required widely - 1

Descriptive Attributes

21. "ENTITY TYPE" Names the class of entities to which this entity belongs, e.g., Data Element. Can apply to all entities.

:Indispensable - 15
:Useful, but low priority - 1
:No rating - 2

:Two interviewees noted that entity type as an attribute is unnecessary in the commercial DDS they use, because it is implicit in the structure of the entry.

22. "KEYWORDS" Words or phrases which categorize an entity, chosen from a list of keywords developed by each organization. They can be used to relate and cross-reference entities. Can apply to all entities.

:Indispensable - 10
:Required widely - 6
:Useful, but low priority - 2

:One interviewee believed that this capability could cause a serious overhead problem.

23. "DESCRIPTION" Defines the purpose and use of the entity. A free-form text field, it is supplemented by the Keywords. Can apply to all entities.

:Indispensable - 17

:Required widely - 1

24. "STORAGE LOCATION" The place of physical storage of documents, files, and other information entities in a manual system. In the case of a distributed or multi DDS-DBMS environment, the computer system, files, libraries, or multiple locations where information entities are stored.

:Indispensable - 3
:Required widely - 2
:Limited application - 5
:Useful, but low priority - 4
:Not required - 2
:No rating - 2

25. "SCHEDULE OF USE" Description of the schedule of production of a report, execution of a system or program, the collection of data, etc.

:Indispensable - 3
:Required widely - 3
:Limited application - 4
:Useful, but low priority - 3
:Not required - 3
:No rating - 2

:One interviewee commented that STORAGE LOCATION and SCHEDULE OF USE were not needed because other software packages perform these functions in her environment.

Security and Standardization Attributes

26. "SECURITY REQUIREMENTS" Documents the requirements that must be met by any user who wishes to access the actual information described by a DDS entity.

:Indispensable - 11
:Required widely - 3
:Limited application - 3
:Useful, but low priority - 1

27. "REFERENCE STANDARD" Names the standard to which the entity conforms, e.g., a Federal Information Processing Standard.

:Indispensable - 5
:Limited application - 3
:Useful, but low priority - 8
:No rating - 2

28. "STANDARDIZATION STATUS" For entities which do not fully conform to a reference standard, this expresses the degree to which standardization has been achieved. Examples could be: "Proposed," "De-facto," etc.

:Indispensable - 7
:Required widely - 2
:Limited application - 1
:Useful, but low priority - 4
:Not required - 2
:No rating - 2

Documentation

These attributes can apply to all entities.

29. "REFERENCE DOCUMENTATION" Provides information about general or special documentation of an entity. It may consist of the title and reference number of documents maintained off-line or in hardcopy form.

:Indispensable - 7
:Required widely - 3
:Limited application - 2
:Useful, but low priority - 4
:Not required - 1
:No rating - 1

:An interviewee thought that this function should be accomplished by a relationship entry, which would relate the entity to the documentation (also described in the DDS).

30. "ENTRY CHANGES" Stores data which summarize the changes made to the DDS entry since its creation. May consist of such parts as: creation date; version effective date; last modification date; last user to modify; and total number of modifications.

:Indispensable - 6
:Required widely - 5
:Limited application - 1
:Useful, but low priority - 4
:Not required - 2

:One interviewee felt that this is a form of version control.

REPRESENTATION AND ENVIRONMENT ATTRIBUTES

Representation Attributes

This section covers attributes that can apply to information entities.

31. "REQUIREMENTS INFORMATION" This is a group of attributes used for requirements forecasting or analysis. It could contain such attributes as: "Growth Rate of Database or File," "Maximum Size of Database or File," "Number of Users."

:Indispensable - 5
:Required widely - 1
:Limited application - 4
:Useful, but low priority - 6
:Not required - 2

32. "FILE CHARACTERISTICS" This group contains such attributes as: "Type of File" (manual or automated), "Medium of File," "Label of File," "Blocking Factor of File."

:Indispensable - 7
:Required widely - 6
:Limited application - 2
:Useful, but low priority - 3

33. "RECORD CHARACTERISTICS" A group of attributes describing the logical format of a record, containing such attributes as: "Arrangement of Data Elements or Data Groups in Record," "Record Key Identifier," "Length of Record."

:Indispensable - 12
:Required widely - 5
:Useful, but low priority - 1

34. "DATA GROUP CHARACTERISTICS" A group of attributes describing a data group. Can contain such attributes as: "Arrangement of Elements in Group."

:Indispensable - 12
:Required widely - 3
:Limited application - 1
:Useful, but low priority - 2

35. "DATA ELEMENT CHARACTERISTICS" A group of attributes describing data element format, which can contain such attributes as: "Format Name," "Character Type," "Length," "Justification," "Sign."

:Indispensable - 15
:Required widely - 3

:One agency had a requirement to maintain, in the DDS, several data element entries with the same name. These are the same entities, but they differ in their representation. There could be different versions, over time, for each entity. They commented that this can be viewed as an entity stored in one representation on the computer and presented externally in another. (In this case, the DDS should cross-reference the conversion subroutines.) Alternately, it can be viewed as an entity that is represented in one format in one information system and in a diffrent format in another information system.

:Another interviewee suggested that data element values for validation purposes be added to the list of attributes. Another suggested fill values, void value, and precision. Fixed vs. variable default values was also raised as an area for consideration.

36. "INPUT/OUTPUT CHARACTERISTICS" Applies to Reports, Forms and Documents. This group contains such attributes as: "Headings," "Arrangement of Elements," "Number of Copies."

:Indispensable - 3
:Required widely - 2
:Limited application - 9
:Useful, but low priority - 3
:No rating - 1

37. "PROGRAM AND SUBPROGRAM CHARACTERISTICS" This group includes such attributes as "Language of Program," "Statement of Software Logic."

:Indispensable - 6
:Required widely - 3
:Limited application - 4
:Useful, but low priority - 5

:One interviewee draws a distinction between PROGRAM (executable code) and SUBPROGRAM (non-executable because the code is incomplete).

Environment Attributes

This section covers attributes that apply to system, subsystem, and program entities.

38. "HARDWARE" Description (manufacturer and model) of mainframe, tape drives, disks, drums, data transmission devices, etc.

:Indispensable - 4
:Required widely - 1
:Limited application - 3
:Useful, but low priority - 7
:Not required - 3

39. "SOFTWARE SUPPORT SYSTEMS" The generalized software that supports the program, system or database, such as the DBMS and support utilities.

:Indispensable - 3
:Required widely - 3
:Limited application - 4
:Useful, but low priority - 6
:Not required - 2

40. "EXECUTION ENVIRONMENT" This includes such attributes as: "Operating Mode (batch, online, other)," "Memory Requirements," etc.

:Indispensable - 5
:Required widely - 1
:Limited application - 4
:Useful, but low priority - 6

:Not required - 2

:Two interviewees who gave HARDWARE, SOFTWARE SUPPORT SYSTEMS, and EXECUTION ENVIRONMENT (#38, 39, 40) a low rating noted that they expected these items to be more important to their environment within the next few years.

Other Attributes

41. "USER-DEFINED ATTRIBUTE (EXTENSIBILITY)" This provides the ability to extend the original range of attributes to include those unique to a given user's environment.

:Indispensable - 8

:Required widely - 4

:Limited application - 3

:Useful, but low priority - 2

:No rating - 1

:One interviewee commented that in the early stages of using a DDS this item can be counter-productive because it could tend to propagate non-standard descriptions.

ADDITIONAL NOTES ON ATTRIBUTES

1. Several interviewees thought that these attributes should be added to the list:

(a) Criticality - for file and system entities, i.e., a way to indicate the importance of data for use in analyzing the impact of system failure on non-ADP program operations, response to the public, decision-making, etc.

(b) Synonym - differing from the Alias capability in defining different views of an entity, i.e., two systems using the same logical piece of information (two similar data elements) with different names and representations. There should be an ability to describe these data elements for purposes of standardization and display of information.

(c) Related Term - to show related information such as city and state.

2. One agency thought that these should be deleted:
 - (a) Reference Standard
 - (b) Standardization Status

II. RELATIONSHIPS

This refers to the capability to define relationships among entities in the DDS in order to access one entity description from a related one, or trace information flow and usage in information systems. The following items do not cover the full complexity or all aspects of possible relationships, but serve to sample the range of Federal agency needs.

42. "COMPONENT relationships" Capability to indicate the component entities of an aggregated entity; such as the data elements contained in a record.

:Indispensable - 17
:No rating - 1

43. "INPUT relationships" Capability to indicate the entities that are input to a program, system, or process, such as the forms and files used in an update procedure.

:Indispensable - 13
:Required widely - 1
:Limited application - 2
:Useful, but low priority - 2

44. "OUTPUT relationships" Capability to indicate the entities provided as output of a program, system, or process.

:Indispensable - 13
:Required widely - 1
:Limited application - 2
:Useful, but low priority - 2

45. "ANCESTRAL relationships" Capability to indicate the derivation of one entity from another, or replacement of an obsolete entity by another entity.

:Indispensable - 3
:Required widely - 3
:Limited application - 6
:Useful, but low priority - 5
:Not required - 1

46. "PROCESS CYCLE relationships" Capability to indicate the flow relationships among programs, subsystems, or tasks in an operating cycle or indicate the sequence of information processing tasks.

:Indispensable - 2
:Required widely - 4
:Limited application - 5
:Useful, but low priority - 7

:In two agencies this function is performed by other software packages. The DDS is not used in conjunction with them.

47. "USER relationships" Capability to indicate the financial and programmatic controllers and users of DDS entities.

:Indispensable - 4
:Required widely - 4
:Limited application - 4
:Useful, but low priority - 3
:Not required - 2
:No rating - 1

48. "USER-DEFINED relationships" Capability to define additional relationship types.

:Indispensable - 8
:Required widely - 2
:Limited application - 5
:Useful, but low priority - 3

III. OPERATING MODES and CONTROL FUNCTIONS of the DDS

A. OPERATING MODES

49. "BATCH MODE" For bulk update or initial loading

:Indispensable - 14
:Required widely - 2
:Limited application - 1
:Useful, but low priority - 1

50. "BATCH MODE" For retrieval or report production

:Indispensable - 12
:Required widely - 4
:Limited application - 1
:Useful, but low priority - 1

51. "INTERACTIVE MODE" For Update/Retrieval

:Indispensable - 13
:Required widely - 3
:Limited application - 1
:Useful, but low priority - 1

B. CONTROL FUNCTIONS

52. "COPY" Creates a new DDS entry with same attributes and relationships as an existing one.

:Indispensable - 12
:Required widely - 3
:Limited application - 1
:Useful, but low priority - 2

DDS SECURITY LEVELS Enforces rules for DDS security and integrity on all commands which are syntactically correct. See items 53, 54, and 55.

53. "DICTIONARY ACCESS PROTECTION" Applies to the contents of the DDS as a whole; determines who may and may not have access to any of the DDS entries.

- :Indispensable - 13
 - :Required widely - 3
 - :Limited application - 2
54. " ENTITY TYPE SECURITY" Rules governing access to all entities of a given type; for instance, some users may have access to data element entries, but not to system entries.
- :Indispensable - 11
 - :Required widely - 3
 - :Limited application - 1
 - :Useful, but low priority - 3
55. "ENTRY OCCURRENCE SECURITY" Rules applying to the access of individual DDS entries. For example, a specific user may have permission to modify only certain data element entries.
- :Indispensable - 15
 - :Required widely - 2
 - :Limited application - 1
56. "DDS INTEGRITY EDIT" Checks rules that protect against inadvertent errors to the DDS contents. For example, a command which would create an "orphan" (an entity without relationship to any other entity) would not be executed.
- :Indispensable - 12
 - :Required widely - 1
 - :Limited application - 1
 - :Useful, but low priority - 1
 - :Not required - 2
 - :No rating - 1
- :One interviewee commented that the ability to enter incomplete entries in building the dictionary would be preferable to a rigid integrity edit. Another thought it was a good idea, but disagreed with the example.
57. "ENTITY/ATTRIBUTE EDIT" Validates the entities, attributes and relationships mentioned in the user command against DDS internal parameters. This can include table-lookup validation of standard codes for certain attributes.

:Indispensable - 10
:Required widely - 7
:Useful, but low priority - 1

IV. LANGUAGE FUNCTIONS AND FORMAT

58. "CREATE" Define and load DDS entries.

:Indispensable - 17
:Required widely - 1

59. "UPDATE" Insert, modify and delete DDS entries.

:Indispensable - 18

TYPES OF RETRIEVAL:

60. "BY ENTITY NAME" Retrieve attributes of an entity.

:Indispensable - 18

61. "BY ATTRIBUTE" Retrieve on attribute value, such as a specific keyword.

:Indispensable - 13
:Required widely - 4
:Useful, but low priority - 1

62. "BY RELATED ENTITY" Retrieve entities based on their relationship to another entity.

:Indispensable - 13
:Required widely - 3
:Limited application - 1
:Useful, but low priority - 1

63. "BY COMPOUND LOGIC" Retrieval based on combinations of specified attributes or relationships.

:Indispensable - 9
:Required widely - 6
:Limited application - 3

64. "RETRIEVAL BY KEYWORD" Information obtained in KWIC (keyword-in-context) or KWOC (keyword-out-of-context) format.

:Indispensable - 13
:Required widely - 2
:Useful, but low priority - 3

TYPES OF COMMANDS:

65. "STRING-ORIENTED COMMANDS" English-like syntax for ad hoc retrieval.

:Indispensable - 10
:Required widely - 5
:Limited applicaton - 2
:Useful, but low priority - 1

66. "DISPLAY-ORIENTED COMMANDS" Use of tabular displays or preformatted screens to guide entry of DDS data or add arguments for commands.

:Indispensible - 5
:Required widely - 8
:Limited application - 5

67. "MENU DRIVEN QUERIES" Cues user as to available options.

:Indispensible - 7
:Required widely - 4
:Limited application - 1
:Useful, but low priority - 6

68. "TUTORIAL" Online teaching capability which instructs users concerning use of the DDS.

:Indispensible - 5
:Required widely - 3
:Limited application - 3
:Useful, but low priority - 6
:Not required - 1

69. "FIXED-FORMAT COMMANDS" Rigid command structures into which users may insert parameters.

:Indispensible - 5

:Required widely - 6
:Limited application - 1
:Useful, but low priority - 4
:Not required - 2

:One interviewee commented that the closer a command is to English, the easier it is for a user to accept.

V. DDS INTERFACES and FUNCTIONS

REPORT WRITER CAPABILITY:

70. "FIXED FORMAT REPORTS"

:Indispensable - 12
:Required widely - 2
:Limited application - 1
:Useful, but low priority - 2
:Not required - 1

71. "PARAMETERIZED REPORTS"

:Indispensable - 9
:Required widely - 5
:Limited application - 2
:Useful, but low priority - 2

:One interviewee thought that this item was unclear and should be renamed.

72. "AD HOC REPORTS"

:Indispensable - 10
:Required widely - 4
:Limited application - 2
:Useful, but low priority - 2

73. "HOST LANGUAGE INTERFACE" Allows an application program to access the contents of the DDS.

:Indispensable - 11
:Required widely - 1
:Limited application - 3

:Useful, but low priority - 2
:Not required - 1

74. "SOURCE LANGUAGE SCAN" Reads computer program source code and generates DDS input.

:Indispensable - 5
:Required widely - 5
:Limited application - 1
:Useful, but low priority - 5
:Not required - 2

:One interviewee commented that this item has little usefulness in practice; a second stated that it is very desirable to use this capacity in building the DDS.

75. "STORED QUERIES" Capability to save common query procedures for later reuse.

:Indispensable - 7
:Required widely - 3
:Limited application - 2
:Useful, but low priority - 5
:Not required - 1

GENERATION METHOD

Supplies source data from the DDS for use by other automated systems, such as a DBMS or application program. The generated source can be delivered in the following ways.

76. "MANUAL REFERENCE" The user must access various DDS entries individually and assemble the source data by a sequence of command actions.

:Indispensable - 3
:Required widely - 2
:Limited application - 3
:Useful, but low priority - 5
:Not required - 5

:One interviewee thought that this and the next item (DIRECT OUTPUT)'s functions should be automated.

77. "DIRECT OUTPUT" The source data is maintained within the DDS and written to an external file.

:Indispensable - 11
:Required widely - 2
:Limited application - 4
:Useful, but low priority - 1

78. "CREATION" The DDS contains sufficient data to allow it to create the required information in a desired format, which may be further edited to add hardware dependent features.

:Indispensable - 9
:Required widely - 6
:Limited application - 1
:Useful, but low priority - 2

79. "INTERFACE TO USER-WRITTEN CODE" Allows insertion of user-written code into the programs which control DDS functions.

:Indispensable - 9
:Required widely - 4
:Limited application - 2
:Useful, but low priority - 3

GENERATION REQUIREMENTS

This is the source data generated for ultimate use by other automated systems. Types of generated source can include:

80. "OPERATING SYSTEM CONTROL LANGUAGE"

:Indispensable - 2
:Required widely - 2
:Limited application - 4
:Useful, but low priority - 6
:Not required - 3
:No rating - 1

81. "DATA DIVISION FOR APPLICATION PROGRAMS" List languages you need or may require in the future.

:Indispensable - 10
:Required widely - 2
:Limited application - 4

:Not required - 2

:The languages listed were:

ALC (1)
Assembler (1)
COBOL (5)
PL/I (4)
FORTRAN (2)
ADA (2)
PASCAL (2)

82. "DBMS CONTROL DATA" e.g., Data Definition Language, PSB's, etc. Please list kinds of control data you may need now or in the future.

:Indispensable - 10
:Required widely - 2
:Limited application - 3
:Useful, but low priority - 2
:Not required - 1

:There was a wide variety of responses to this item:

DDL and DML
Schemas and subschemas
Control mechanisms, such as PSB, PCB, APB, and DBD

83. "TEST DATA"

:Indispensable - 5
:Required widely - 2
:Limited application - 3
:Useful, but low priority - 6
:Not required - 2

84. "EDIT/VALIDATION TABLES"

:Indispensable - 11
:Required widely - 1
:Limited application - 4
:Useful, but low priority - 2

VI. STANDARD DDS PRINTED OUTPUTS

85. "CATALOG REPORTS" These reports contain the identifier and all attribute values of all entities.

:Indispensable - 16
:Required widely - 1
:Not required - 1

:One interviewee felt that this report should also include relationships.

86. "ENTITY RETRIEVAL REPORTS" As above, but only some entities are described. The user specifies the attribute values of the entities of interest.

:Indispensable - 14
:Required widely - 2
:Limited application - 1
:Useful, but low priority - 1

87. "CROSS-REFERENCE REPORTS" These reports display the attributes of entities which have relationships specified by the user. Example: Display identifiers of all reports and their related programs which contain both of the elements "Employee ID" and "Salary."

:Indispensable - 14
:Required widely - 3
:Limited application - 1

88. "VERSION CONTROL REPORT" A display of all versions of an entity, noting status differences and dates of creation, copied from/to, etc.

:Indispensable - 6
:Required widely - 4
:Limited application - 3
:Useful, but low priority - 3
:Not required - 2

89. "INTEGRITY CONTROL REPORT" Exception reports for entity occurrences that do not satisfy the integrity controls. An additional report displays tables or parameters controlling this function.

:Indispensable - 9

- :Required widely - 3
 - :Limited application - 2
 - :Useful, but low priority - 3
 - :Not required - 1
90. "EXTENSIBILITY REPORT" A description of all user-defined entities, attributes, and relationships including date of creation, extent of use, etc.
- :Indispensable - 7
 - :Required widely - 3
 - :Limited application - 4
 - :Useful, but low priority - 3
 - :Not required - 1
91. "SECURITY REPORT" A display of tables or parameters controlling the security functions.
- :Indispensable - 11
 - :Required widely - 1
 - :Limited application - 3
 - :Useful, but low priority - 3
- :One interviewee noted that this report should be restricted to DBA use. Another thought that any security reports must also be controlled by security procedures.
92. "THESAURUS-TYPE REPORT" A display of all keywords allowable.
- :Indispensable - 9
 - :Required widely - 4
 - :Limited application - 1
 - :Useful, but low priority - 4
93. "USAGE STATISTICS REPORT" Reports on DDS and data usage statistics.
- :Indispensable - 4
 - :Required widely - 3
 - :Limited application - 7
 - :Useful, but low priority - 4

ADDITIONAL COMMENT ON REPORTS

One interviewee stated that there should be a differentiation between reports from batch runs and outputs from on-line queries.

VII. DDS ADMINISTRATOR SUPPORT and UTILITY FUNCTIONS

94. "BACKUP/RECOVERY" Selective or complete dump of DDS content in a format to facilitate complete or partial reload to recover from various failures.

:Indispensable - 15
:Required widely - 1
:Useful, but low priority - 2

95. "REORGANIZE" Utilities to facilitate reorganizing DDS contents after a period of usage and updating; write entire DDS contents onto tape in serial format, translate and reformat, and then reload.

:Indispensable - 14
:Required widely - 3
:Useful, but low priority - 1

96. "LOAD/RELOAD" Reformats the contents of the DDS into a form suitable for transferring, by tape or other medium, to another DDS on the same or another computer system, then reverses the process on the new system.

:Indispensable - 14
:Required widely - 3
:Not required - 1

5. CONCLUSIONS

Preliminary conclusions on Federal agency requirements for the FIPS DDS are included in this chapter. Technical issues and items which require further analysis are highlighted. These issues and unresolved items will be studied in more depth during the next year. These conclusions and issues are presented to stimulate further discussion with Federal agencies, DDS vendors, and others knowledgeable about the implementation and use of data dictionaries. Conclusions, and proposed solutions, modified as appropriate by further discussions, will serve as input to the development of the FIPS DDS Functional Specifications.

5.1 Specific Conclusions and Unresolved Issues

Specific conclusions about Federal requirements and issues that require further analysis are based primarily on ratings assigned to items in the Rating Form. The conclusions are substantiated by the interview question responses and many of them are supported by comments received on the Prospectus.

Due to the small size of the sample (eighteen Rating Forms), sophisticated statistical techniques are not necessary to analyze the results. Instead, the number of responses for "indispensable" and "required widely" were simply added together to obtain a categorization of the features:

- o Group 1 contains those items which received 14-18 "indispensable" and "required widely" ratings. Group 1 items are considered as strong candidates for inclusion in the FIPS DDS.
- o Group 2 contains items which received a total of 9-13. These items are candidates for possible inclusion.
- o Group 3 contains items that received a wide variation in ratings, but fewer than 9 "indispensable" and "widely required" ratings. For example, one item received 8 "indispensable" and "widely required" and 8 "useful, but low priority" ratings. No rating was given in the remaining two instances. The status of these items is unresolved pending further analysis and additional review with Federal agencies.

- o Group 4 contains items where 11 or more of the ratings were "limited application," "useful, but low priority" or "not required." The preliminary conclusion is that these items will not be required.

Results from this categorization are organized below in the same manner as in the Rating Form, i.e., Entities, Attributes, and Relationships are presented first.

5.1.1 DDS Entities.

Group 1 - Strong Candidates for FIPS DDS

Data Element
Data Group
Record
File
Database
Report
System
Program

Group 2 - Possible Candidates for FIPS DDS

Schema
Form
Program Subsystem
User
· User-defined (Extensibility)

Group 3 - Unresolved

Document
Manual Process

Group 4 - Unlikely Candidate for FIPS DDS

No entities appear in this group.

5.1.2 DDS Attributes.

Group 1 - Strong Candidates for FIPS DDS

DDS Name
Version Identifier
Status
Aliases
Entity Type
Keywords
Description
Security Requirements
Record Characteristics
Data Group Characteristics
Data Element Characteristics

Group 2 - Possible Candidates for FIPS DDS

- Identification Symbol
- Reference Documentation
- Entry Changes
- File Characteristics
- Program and Subprogram Characteristics
- User-defined (Extensibility)
- Standardization Status

Group 3 - Unresolved

- Storage Location
- Schedule of Use
- Reference Standard
- Software Support Systems

Group 4 - Unlikely Candidate for FIPS DDS

- Requirements Information
- Input/Output Characteristics
- Hardware
- Execution Environment

5.1.3 Relationships.

Group 1 - Strong Candidates for FIPS DDS

- Component
- Input
- Output

Group 2 - Possible Candidates for FIPS DDS

- User-defined (Extensibility)

Group 3 - Unresolved

- Ancestral
- User

Group 4 - Unlikely Candidate for FIPS DDS

- Process Cycle

The entities, attributes, and relationships which appear in Group 1 support the consensus view that the most important use of a DDS is and will continue to be as a tool to inventory, describe, and standardize data. There was confusion, however, among three of the entities: Report, which appears in Group 1; Form, which is in Group 2; and Document, which appears in Group 3. These need to be reviewed, clarified, and possibly combined.

Entity Extensibility, Attribute Extensibility, and Relationship Extensibility all appear in Group 2. This implies that capabilities such as extensibility which support flexibility of use should be supported as an optional

capability.

Although several agencies cited a strong requirement to document data processed manually, the Manual Process entity appears in Group 3 as unresolved. There was a rating variation of 7 who felt that it was "indispensable" or "widely required" and 7 who felt it was "not required" or "useful, but low priority."

Inconsistent results were obtained for the Program and Program Subsystem entities. Program Subsystem, which is defined as a collection of related programs, appears in Group 2. Program, which is more narrowly defined as a subprogram or module, is in Group 1. The attributes Program and Subprogram Characteristics, which support these two entities, appear in Group 2. Although Standardization Status, Requirements Information, Input/Output Characteristics, Hardware, and Execution Environment attributes and the Process Cycle relationship are shown as unlikely candidates, further review may be warranted. Each of these items received 5 or more "indispensable" or "required widely" ratings.

5.1.4 Operating Modes and Control Functions.

5.1.5 Operating Modes.

Group 1 - Strong Candidates for FIPS DDS

Batch Mode for Update or Loading

Batch Mode for Retrieval or Report Production

Interactive Mode for Update/Retrieval

Group 2, 3, and 4

None of the Operating Modes appear in these groups.

5.1.6 Control Functions.

Group 1 - Strong Candidates for FIPS DDS

Copy

Security Edit - Dictionary Access Protection

Security Edit - Entity Type Security

Security Edit - Entity Occurrence Security

Entity/Attribute Edit

Group 2 - Possible Candidates for FIPS DDS

DDS Integrity Edit

Group 3 and 4

None of the Control Functions appear in these groups.

Batch and interactive operating modes and all except one control function received ratings which indicate that they are strong candidates for the FIPS DDS. The three types of Security Edits, however, need more investigation because their ratings reflect a much greater concern for security, particularly at the entity occurrence level, than was expressed during the actual interviews. The DDS Integrity Edit appears as a possible candidate. Several agencies felt that stringent integrity edits might prohibit entry of incomplete entities when complete description and definitions were not available. This concern was expressed by some low ratings which offset the 12 "indispensable" and 1 "widely required" ratings.

5.1.7 Language Functions and Format.

Group 1 - Strong Candidates for FIPS DDS

- Create
- Update
- Retrieval by Entity Name
- Retrieval by Attribute
- Retrieval by Related Entity
- Retrieval by Compound Logic
- Retrieval by Keyword
- String-Oriented Commands

Group 2 - Possible Candidates for FIPS DDS

- Display-Oriented Commands
- Menu-Driven Queries
- Fixed-Format Commands

Group 3 - Unresolved

- Tutorial Commands

Group 4 - Unlikely Candidate for FIPS DDS

No Language Functions appear in this group.

All create, update, and retrieval capabilities are widely required. Although Retrieval by Keyword received ratings high enough to place it in Group 1, several interviewees cautioned that they had encountered serious overhead problems using it. It may be preferable, after implementation considerations are investigated in greater depth, to include this as an optional capability.

Language formats appear in all three groups. Most of the possible formats are in Group 2, as possible candidates for the FIPS DDS. Only one, String-Oriented Commands, is in Group 1. Tutorial Commands received a wide variation in ratings, which resulted in its placement in Group 3. Some agencies feel that Tutorial Commands are necessary to support the major ease-of-use requirement. Other agencies, which have had more experience in using a DDS, feel that

this capability has "limited application" or is a "low priority."

Agencies were asked if they felt it important for the DDS and DBMS language to be similar. There was no consensus -- some feel they should be similar, but others do not think this is a requirement. A DDS vendor, commenting on the Prospectus, recommended that the FIPS DDS specify language functions and capabilities, but not the precise syntax. There is widespread interest on the subject of DDS language(s), but no definitive results on Federal requirements have been obtained. This is an issue that needs further analysis and ultimate agency resolution to answer such questions as:

- What human factors need to be considered for both ADP and non-ADP users?
- Is a single, integrated user language for DDS input and operating control feasible?
- Should the FIPS DDS specify more than one interactive language? What should be the capabilities and structure of the interactive language(s)?
- Is CODASYL compatibility desirable?

5.1.8 DDS Interfaces and Functions.

Group 1 - Strong Candidates for FIPS DDS

Fixed Format Reports

Parameterized Reports

Ad Hoc Reports

Generation Method:

(Means by which source data is supplied from the DDS to other systems)

Creation - DDS creates data in format desired

Group 2 - Possible Candidates for FIPS DDS

Host Language Interface

Source Language Scan

Stored Queries - DDS can save common queries for reuse

Generation Method:

Direct Output - source data maintained within DDS

Interface to User-Written Code

Generation Requirements:

(Type of source data generated for use by other systems)

Data Division for Application Programs
DBMS Control Data

Edit/Validation Tables

Group 3 - Unresolved

No DDS Interface appears in this group.

Group 4 - Unlikely Candidate for FIPS DDS

Generation Method:

Manual Reference - user must assemble
source data

Generation Requirements:

Operating System Control Language
Test Data

Clear results were not obtained on the requirement to generate source data for ultimate use by other automated systems. Most of the types of source data that Federal agencies want the DDS to generate appear in Group 2 as possible candidates. The preferred generation method is for the DDS to create data descriptions in the format required by other software. Respondents definitely do not want to access various DDS entries individually and assemble source data through a sequence of command actions.

Although DBMS Control Data appears as a possible candidate, twelve out of eighteen rated it as "indispensable" or "widely required." In many of the agencies, a DDS-DBMS interface is fundamental to control of data. There is currently no FIPS DBMS, and at least three data models are used as the basis for current systems. Respondents listed as requirements many different types of control data for different DBMS's. It must still be determined, with these conditions, what scope and degree of DDS-DBMS interaction should be specified in the FIPS DDS.

Although seven respondents rated the generation of Test Data as "indispensable" or "widely required," the majority do not consider this capability a requirement. Only four out of eighteen rated Operating System Control Language generation as a requirement, which indicates that this item should be excluded from further review.

All interfaces (Host Language Interfaces, Source Language Scan, and Interfaces to User-Written Code) received ratings which make them possible candidates for the FIPS DDS. Further analysis is required to assess what impacts these interfaces may have on other software, including operating system services, performance, etc.

5.1.9 Standard Printed Outputs.

Group 1 - Strong Candidates for FIPS DDS

Catalog Reports
Entity Retrieval Reports
Cross-Reference Reports

Group 2 - Possible Candidates for FIPS DDS

Version Control Report
Integrity Control Report
Extensibility Report
Security Report
Thesaurus-Type Report

Group 3 - Unresolved

None of the Printed Outputs appear in this group.

Group 4 - Unlikely Candidate for FIPS DDS

Usage Statistics Report

Ratings results show that a report writer capability to produce Fixed Format, Parameterized, and Ad-hoc Reports is desirable. In responding to questions about current DDS use, many interviewees emphasized that a report customization capability is needed. This need is strongest in those agencies where the DDS reports are being used widely by non-ADP users.

The placement of the Integrity Control and Extensibility Reports in Group 2 as possible candidates is consistent with the DDS Integrity Edit ratings and the Entity Extensibility, Attribute Extensibility and Relationship Extensibility ratings. Inconsistent results, however, were obtained on the Version Control, Security, and Thesaurus-Type Reports and the DDS features that support these reports. The Version Identifier attribute and the three types of Security Edits received ratings that put them in Group 1, likely FIPS DDS candidates. Likewise, the Keyword attribute appears in Group 1. The associated Thesaurus-Type Report is in Group 2, a possible candidate. This is another area that needs further study and ultimate agency resolution.

5.1.10 DDS Administrator Support and Utility Functions.

Group 1 - Strong Candidates for FIPS DDS

Backup/Recovery
Reorganize
Load/Reload

Group 2, 3, and 4

No Functions appear in this group.

As can be seen, all three of the DDS support and utility functions received ratings that make them likely FIPS DDS candidates. Further analysis, however, is necessary to determine how the DDS contents should be unloaded for reorganization or transport purposes.

Two issues were identified during the interviews which are not reflected in the rating results. These are:

- o DDS implementation on small computers (minis and micros). There was no consensus on the need for a DDS that would operate on small computers. If this is a requirement, the basic capability of the FIPS DDS may be dictated by what is technically and economically feasible to implement in such an environment.
- o Features required to support distributed database systems and distributed processing systems. Several questions are associated with this issue. How and to what extent should the FIPS DDS support distributed processing applications? Are any unique features, entities, and attributes needed? Is a special interface to network control programs needed?

5.2 General Conclusions

Two levels of DDS use in the Federal Government were identified. There was a consensus that the most important use of a DDS is and will continue to be oriented toward data management, e.g., to inventory, describe, and standardize data. Some of the agencies interviewed use the DDS solely for these purposes, without regard to any ADP application. Most agencies, however, also use the DDS to help manage their ADP resources. At this level, the DDS assists agencies in their ADP system planning, requirements analysis, change-impact analysis and documentation. The DDS also contributes to standardizing the use of data in ADP systems when it is used to generate data descriptions for the DBMS and application programs.

Flexibility and ease-of-use are the two major Federal requirements for both levels. Federal agencies want to use the FIPS DDS in a variety of hardware and software environments. Many agencies consider it highly desirable to have a DDS that can interface with a variety of DBMS's and programming languages.

To facilitate ease-of-use, particularly use of DDS printed outputs by non-ADP users, an output customization capability is desirable. There is also a need to support ease-of-use in DDS input and update.

APPENDIX A : QUESTIONNAIRE

FEDERAL AGENCY INTERVIEW GUIDE

I. BACKGROUND INFORMATION

1. Date of interview:
2. Interviewers:
3. Name of Agency and Subdivision(s) Interviewed:
4. Name, Title and Type of Involvement with a Data Dictionary for Each Person(s) Interviewed:

II. QUESTIONS ABOUT NBS PROJECT AND RATING FORM

1. Do you have any questions about the project as a whole or about this interview?

2. Because terms and their meanings differ from place to place, the working definition of DDS which we will use in this interview is:

A Data Dictionary System (DDS) is a computer software system that provides for recording, storing, and processing information about an organization's significant data and data processing entities.

Do we share a common definition of DDS, or if we differ, please identify the specific points where we differ.

3. Is the terminology used in the Rating Form clear and understandable?
If not, what terms did you find unclear or subject to multiple interpretation?

4. In completing the Item Rating Form, did you have any questions about the technical or economic feasibility of an item?

If so, please explain.

5. Are there any items missing from the Rating Form that you believe should be included in the eventual standard?

If so, what entities, attributes, relationships and capabilities should be included?

6. Should a capability for system- or hardware-dependent features be included?

7. Should physical storage attributes be included?

8. Were there any items included that you feel should be omitted from the eventual standard?

If so, please explain.

III. ENVIRONMENT OF DATA DICTIONARY SYSTEM

1. What Data Dictionary System(s) (DDS) is your organization using?

2. What is the purpose of your DDS(s)?

3. If you have more than one DDS, do they communicate with each other?

4. Please describe the environment of your DDS; that is, not your entire operating environment, but the portion of it with which the DDS interacts directly. Please touch on:
 - a. Mainframe(s) on which the DDS runs: manufacturer, model, size of memory.

 - b. Operating system and mode?

 - c. Is the mainframe part of a network?
If so: Are minicomputers or microcomputers included?

- d. What role does the DDS have in the network?
5. How much of your computer resources does your current DDS consume in regard to such things as memory and on-line storage?
- a. Is this an acceptable level of system overhead?
- b. Is the response time acceptable?
6. How much data is stored in your data dictionary, e.g., approximately how many entity types and occurrences?
7. Do you use all the entity types provided by your system?
Please indicate which ones you use.
8. Does the DDS support a few particular applications or is it oriented toward organization-wide data?

9. How is your organization's data - not the dictionary data - organized? (e.g., in files, in databases, etc.)

10. Does your organization use a Database Management System?

If yes:

a. Is more than one DBMS used?

b. Please give the name(s) and vendor(s).

c. What is the relationship of the DDS to the DBMS(s)?

d. Is the DDS definition language similar to the DBMS data definition language?

Do you feel it is important for them to be similar?

11. What programming languages are used in your organization?

12. Which of the programming languages are currently used or may be used in the future with your DDS?

IV. SELECTION OR DEVELOPMENT OF EXISTING DATA DICTIONARY SYSTEM(S)

1. How large a staff was involved in developing or installing your DDS?

2. Describe the overall cost of bringing the system into operation.

3. What is the current level of effort involved in maintaining and enhancing the system?

4. Obtained Commercially:
What were the most important reasons that determined your selection of this particular package?
 - a. Only one available for my computer system.

 - b. Offered by my DBMS vendor.

 - c. Offered by my hardware vendor.

 - d. Recommended by someone I trust.

 - e. Least expensive.

 - f. Had the features I needed. (Which ones were the most important?)

g. Already in place when I came to work for this company/agency.

h. Other. Please be specific.

5. What other packages were evaluated?

6. Obtained from Another Organization or Developed Internally:

If it is not a commercial package, how was it acquired?

a. From another organization.

b. Wrote our own.

c. Other. Please be specific.

7. If it was acquired from another organization, was it necessary to modify it substantially?

If so, what changes did you make?

8. If you developed your own, why was that decision made?

a. No other options available.

b. Packages too expensive.

c. Packages did not meet our needs.

d. Easier than adapting software from other organizations.

e. Other. Please be specific.

V. USES OF CURRENT DATA DICTIONARY

1. Why did your agency obtain (or develop) a DDS?

2. In what ways is the DDS used?

Is it used:

- a. To assist in requirements analysis.

- b. To assist in the design of application software and databases.

- c. To prepare documentation for: application programs, logical database design, physical database design; others.

- d. To generate data descriptions, test data, other.

- e. To control operations and scheduling.

- f. To assist in (or enforce) data and documentation standards.

- g. Other - please be specific.
3. Does your DDS have internal security and integrity controls?
If yes, describe the type of control available.
4. Does the DDS exert control over other systems?
If yes:
 a. What types of systems?
- b. How and for what purpose is the control exerted?
(e.g., supply source data descriptions, collect usage statistics, etc.)
5. Can the DDS be called directly by other systems or can it call other systems without human intervention?
6. Who are the DDS users?
 a. Is ADP expertise required to use the DDS?
 b. Are any of the DDS outputs used by non-ADP users?

If so, for what purpose?

c. What is the approximate total number of DDS users?

7. Who (organizationally) is responsible for your dictionary's:

a. Selection or development?

b. Software installation?

c. Software modification?

d. Day-to-day management?

e. Data input?

f. Report generation?

g. Contents determination?

h. Other - please be specific.

8. What are the shortcomings or disadvantages of your DDS?
- a. Ease-of-use by ADP personnel and end users (non-ADP personnel).
 - b. Scope of entities and attributes that can be described.
 - c. Processing functions, utilities, and overhead.
 - d. Report and query capabilities.
 - e. Other - please be specific.
9. Would these shortcomings be alleviated by items identified in the Rating Form?
If yes, which ones?
10. Do you feel that the items you need have been completely identified through previous questions?
11. What benefits have been obtained by using the DDS?

12. Has the DDS provided all of the benefits you expected in your original planning?

If no, please elaborate.

13. Have you ever converted from one DDS to another?

If yes, what problems did you encounter?

VI. FUTURE PLANS AND ENVISIONED TRENDS

1. Do you have any plans to purchase or obtain new data dictionary software?

If yes:

a. Why?

b. How will the new data dictionary be used?

c. Will you continue to use your existing data dictionary?

If so, how will it interact with the new one?

2. What trends do you envision in the use of data dictionaries in your agency and in general?

3. Do you envision the future use of multiple DDS or one centralized DDS?

4. Do you think a DDS and DBMS should be independent of each other or closely related?

5. Do you envision a future need for a DDS which deals with distributed processing and storage?
If yes, please expand.

6. Is the scope and complexity of DDS software a potential problem for you?

7. Do you think your agency will require a DDS that could run on a minicomputer or other very small computers?

8. How will your organization work with the Federal Information Locator System (FILS) being developed by OMB?

9. Do you feel that guidelines on DDS usage are more important than a software standard?

APPENDIX B: AGENCIES AND SUBCOMPONENTS

The names of the agencies and subcomponents where interviews were conducted are listed in alphabetical order.

Department of Defense

Office of the Assistant Secretary of Defense
(Comptroller)

Defense Communications Agency

Headquarters, Defense Communications Agency

 Staff Database Communications Officer

Defense Communications Engineering Command (DCEC)

National Communication Systems/Defense Communication
 Agency Operation Center (DCAOC)

Department of the Air Force

Headquarters, U.S.A.F. (Air Staff)
 Air Force Data Services Center

Department of the Navy

Naval Data Automation Command
 Navy Regional Data Automation Center,
 Washington, D.C.
 Systems Standards Division

Environmental Protection Agency

Office of Planning and Management
 Management Information and Data Systems Division

General Services Administration

National Archives and Records Service
 Office of Records and Information Management

Department of Health and Human Services

Social Security Administration
 Office of Systems Planning and Control
 Division of User Services

Office of Systems Development

 Division of Management and Technical Support

Department of the Interior

Geological Survey

Office of the Assistant Director for Research

Department of Labor

Office of the Assistant Secretary for Information
and Management

Directorate of Information Technology

Office of Policy and Standards

Employment and Training Administration

Office of Administration and Management

Office of Management and Information Systems

Library of Congress

Automated Systems Office

Small Business Administration

Office of Data Management

Department of the Treasury

Internal Revenue Service

Systems Software and Standards Branch

Veterans Administration

Office of Data Management and Telecommunications

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)		1. PUBLICATION OR REPORT NO.	2. Performing Organ. Report No.	3. Publication Date
		NBSIR81-2354		September 1981
4. TITLE AND SUBTITLE				
<p><i>Federal Requirements for a Federal Information Processing Standard Data Dictionary System</i></p>				
5. AUTHOR(S)				
<p><i>Patricia A. Konig, Judith J. Newton, Roy G. Saltman</i></p>				
6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)		<p>7. Contract/Grant No.</p> NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234		
		<p>8. Type of Report & Period Covered</p>		
9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)				
10. SUPPLEMENTARY NOTES				
<p><input type="checkbox"/> Document describes a computer program; SF-185, FIPS Software Summary, is attached.</p>				
<p>11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)</p> <p><i>This report presents information and preliminary conclusions about Federal agencies' requirements for a Federal Information Processing Standard Data Dictionary System. Some initial requirements were identified through analysis of comments made on the "Prospectus for Data Dictionary System Standard" which describes NBS' efforts to develop a standard. Most of the data used to develop preliminary conclusions on Federal requirements was collected during interviews with Federal Government users and developers of Data Dictionary Systems. Comments received on the Prospectus and data collected during the interviews are summarized. Preliminary conclusions and issues being investigated also are presented.</i></p>				
12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)				
<p><i>Computer program; data dictionary system; data inventory; data management; data standards; database; database management system; documentation; Federal Information Processing Standards Publication; requirements; software.</i></p>				
13. AVAILABILITY		<p>14. NO. OF PRINTED PAGES</p> <p><input checked="" type="checkbox"/> Unlimited</p> <p><input type="checkbox"/> For Official Distribution. Do Not Release to NTIS</p> <p><input type="checkbox"/> Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.</p> <p><input checked="" type="checkbox"/> Order From National Technical Information Service (NTIS), Springfield, VA. 22161</p>		
		<p>80</p> <p>15. Price</p> <p>\$9.50</p>		

